

EXPANDED COG CRITICALITY VALIDATION SUITE

David P. Heinrichs and Soon S. Kim

Lawrence Livermore National Laboratory
P.O Box 808, Livermore, CA 94550
heinrichs1@llnl.gov, kim53@llnl.gov

ABSTRACT

COG is a three-dimensional, continuous-energy, Monte Carlo code, developed by Lawrence Livermore National Laboratory. An established suite of criticality benchmark cases has been formally expanded from 591 to 2,255 intended to comprehensively cover the entire range from thermal to fast neutron spectra under a variety of reflector and moderator conditions, and fissile material types. COG results with ENDF/B-VII.1 and ENDF/B-VIII.0 cross section data have been compared with benchmark values from the International Criticality Safety Benchmark Evaluation Project Handbook. These COG results have been also compared to available MCNP results for the 1,081-case ‘Trkov’ and 119-case ‘Mosteller’ validation suites. Most cases agree within $\pm 3\sigma$ of the benchmark values. About 13% of the total cases are outside of this 6σ range. Possible sources of errors include 1) errors in the cross section data libraries, 2) errors in the modeling the benchmark experiments, or 3) errors in the benchmark experiment itself or its evaluated biases and uncertainties. Good agreement was observed between COG and MCNP results. A major intercomparison project between COG, MCNP, MORET, and SCALE for ENDF/B-VIII.0 and JEFF-3.3 is also in progress. We anticipate that LLNL participation in this project will result in development of significantly more COG benchmark cases as our goal is to overlap the VALID, WHISPER, and IRSN compendia of criticality benchmarks to the extent possible, which will be beneficial to international COG, MCNP, MORET, and SCALE user communities.

KEYWORDS: Monte Carlo, Validation, Benchmark, COG11, Criticality Safety

1. INTRODUCTION

COG [1] is a three-dimensional, continuous-energy, Monte Carlo code, developed by Lawrence Livermore National Laboratory (LLNL). Since 1980, it has been applied to many particle transport applications including nuclear criticality safety, radiation shielding, and reactor physics. Areas of interest at LLNL for nuclear criticality safety include ^{239}Pu , ^{235}U , and ^{233}U systems; however, increased demand from users for the application of COG to other systems motivated LLNL to formally expand the existing validation suite to all major categories of the International Criticality Safety Benchmark Evaluation Project (ICSBEP) Handbook [2]. In this paper, we present COG benchmark results for the expanded LLNL criticality validation suite.

2. EXPANDED CATEGORIES

All benchmark cases were modeled independently by researchers from LLNL, based on information that was taken from Section 3 of the individual evaluations [2]. The number of LLNL benchmark cases (143

Pu, 358 ^{235}U , and 90 ^{233}U) have now been expanded to Pu, HEU (Highly Enriched Uranium), IEU (Intermediate Enriched Uranium), LEU (Low Enriched Uranium), ^{233}U , and mixed cases, reaching a total number of 2,255 cases. The number of benchmark cases in each of these six major categories is summarized in Table 1.

Table 1. Number of Benchmark Cases.

Category	Number of Cases
Pu	570
HEU	817
IEU	188
LEU	363
^{233}U	193
Mixed	124
Total	2,255

These benchmark cases are further categorized by spectra (i.e., fast, intermediate, thermal) and fissionable material form (e.g., metal, compound, solution) as provided in Table A.1, which span the entire range from thermal to fast neutron spectra for a wide variety of fissionable material forms in a variety of reflector and moderator conditions. Calculations for ENDF/B-VII.1 were performed using COG11.1 with: (a) continuous-energy cross sections based on ENDF/B-VII.1 nuclear data, as processed by the International Atomic Energy Agency (IAEA) [3]; (b) probability tables for the unresolved resonance region as processed by Brookhaven National Laboratory using NJOY within the ADVANCE system [4]; and (c) thermal scattering laws using algorithms developed by LLNL [5]. The most recent version, COG11.3, was used for all the benchmark cases with ENDF/B-VIII.0.

3. RESULTS

COG11 results for the 2,255 cases are compared with the benchmark values and their corresponding (1σ) uncertainties as shown in Tables 2 and 3. Most of the 2,255 cases agree with the benchmark values and uncertainties within $\pm 3\sigma$. 298 cases with ENDF/B-VII.1 data and 255 with ENDF/B-VIII.0 data exceed the 6σ range. In general, COG results with ENDF/B-VIII.0 show improved agreement with the benchmark values. Possible sources of discrepant results come from: (a) ENDF/B-VII.1 or ENDF/B-VIII.0 nuclear data; (b) additional errors associated with processing the cross section data; (c) errors in modeling the benchmarks; and (d) errors in the experimental benchmark measurements themselves and their evaluated biases and uncertainties. Cases which lie outside the 6σ range are being intercompared with other researchers using independent methods in a rigorous attempt to determine the causes of such major discrepancies.

**Table 2. COG11.1 Results Compared
with Benchmark Values for ENDF/B-VII.1.**

SD	Pu	HEU	IEU	LEU	²³³U	Mixed
< 1σ	248	503	134	167	120	73
1σ - 2σ	147	142	26	79	38	22
2σ - 3σ	106	64	7	60	9	12
> 3σ	69	108	21	57	26	17
Total	570	817	188	363	193	124

**Table 3. COG11.3 Results Compared
with Benchmark Values for ENDF/B-VIII.0.**

SD	Pu	HEU	IEU	LEU	²³³U	Mixed
< 1σ	406	521	145	169	122	83
1σ - 2σ	91	143	33	84	38	18
2σ - 3σ	32	52	0	51	8	4
> 3σ	41	101	10	59	25	19
Total	570	817	188	363	193	124

As an overall code performance indicator, cumulative χ^2 values are often used as a tool for comparing code results. The χ^2 value is an effective indicator in determining the degree of difference between the calculated and the benchmark values. Here, the cumulative χ^2 is defined as:

$$\chi^2 = \sum_{i=1}^N \frac{(K_{c,i} - K_{b,i})^2}{K_{b,i}} \quad (1)$$

where $K_{c,i}$ and $K_{b,i}$ are the calculated and benchmark k_{eff} values, respectively. Fig. 1 shows cumulative χ^2 values for all six categories of the expanded COG criticality validation suite.

In Fig. 1, ‘71’ in the parentheses represents ENDF/B-VII.1, and ‘80’ represents ENDF/B-VIII.0, respectively. It is observed that the HEU cases performed significantly better than the Pu cases. The poor performance in the PU cases is largely due to the PU-COMP-MIX (Hanford poly-block) cases, which results in the initial steep slope in the Pu curve shown in Fig. 1. The benchmark uncertainties in these cases may be under-estimated, or there may be additional sources of experimental uncertainty not considered in the evaluation. The ²³³U cases also show a steep slope, which raises questions as to the quality of many of the solution experiments that date back to the 1950s. It was further observed that the IEU cases performed better than the ²³³U cases. COG performance with ENDF/B-VIII.0 for HEU was almost the same as ENDF/B-VII.1; however, PU (80) performed much better than PU (71). The performance for LEU (80) is almost the same as LEU (71).

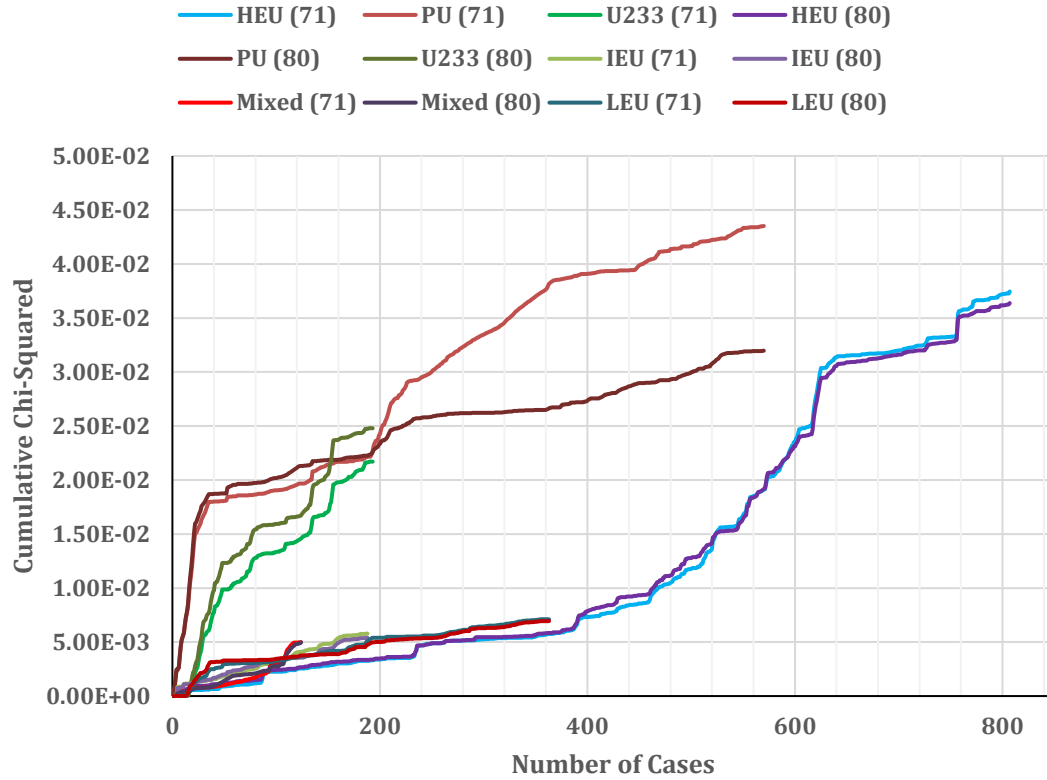


Figure 1. Cumulative χ^2 Values for 2,255-Case COG Results.

3.1 COMPARISON WITH THE IAEA ‘TRKOV’ VALIDATION SUITE

MCNP 6.1 results for 1,081 benchmark cases were published by IAEA using cross sections based on ENDF/B-VII.1 and ENDF/B-VIII.0 nuclear data, processed by researchers from Los Alamos National Laboratory using NJOY [6-7]. COG results were compared to the MCNP results for the 760 matching cases as summarized in Table 4.

Table 4. Number of Benchmark Cases Used for Code Result Comparison.

Category	Number of Cases
Pu	107
HEU	242
IEU	39
LEU	322
²³³ U	32
Mixed	18
Total	760

To test the performance of each set of benchmark cases, the root mean squared error (RMSE) and averaged χ^2 are calculated. The RMSE are defined as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (K_{c,i} - K_{b,i})^2}{N}} \quad (2)$$

where N is the total number of cases. This represents a sample standard deviation of the differences between calculated and benchmark values.

Also used is the averaged χ^2 defined as:

$$\text{Averaged } \chi^2 = \frac{1}{N} \sum_{i=1}^N \frac{(K_{c,i} - K_{b,i})^2}{K_{b,i}} \quad (3)$$

The RMSE for six different sets of results are compared in Fig. 2. COG (71) and COG (80) represent COG results using ENDF/B-VII.1 and ENDF/B-VIII.0, respectively. Likewise, MCNP (71) and MCNP (80) represent MCNP results run by IAEA using ENDF/B-VII.1 and ENDF/B-VIII.0, respectively. Fig. 3 compares the averaged χ^2 for the six subcategories, which shows similar trends as Fig. 2 but with more drastic changes in magnitude.

COG (71) results, shown in blue, are slightly better than MCNP (71) for the Pu, IEU, and ^{233}U cases. The reverse is observed for HEU, LEU, and Mixed cases. The worst performance observed for Mixed cases is due to just one case, namely MMF008-7, a ZEBRA core k-infinity measurement benchmark [8].

Unanticipated discrepancies in probability tables due to differences in processing the ENDF/B-VII.1 data by BNL and LLNL as implemented in COG caused this surprisingly significant difference. As shown in Fig. 2, this problem was resolved when the new probability tables with ENDF/B-VIII.0 data are applied to COG11.3.

COG (80) performed better than COG (71) for Pu, IEU, LEU, and Mixed cases. COG (80) also performed better than MCNP (80) for Pu, IEU, and U233 cases. In general, COG and MCNP performed better with ENDF/B-VIII.0 data library.

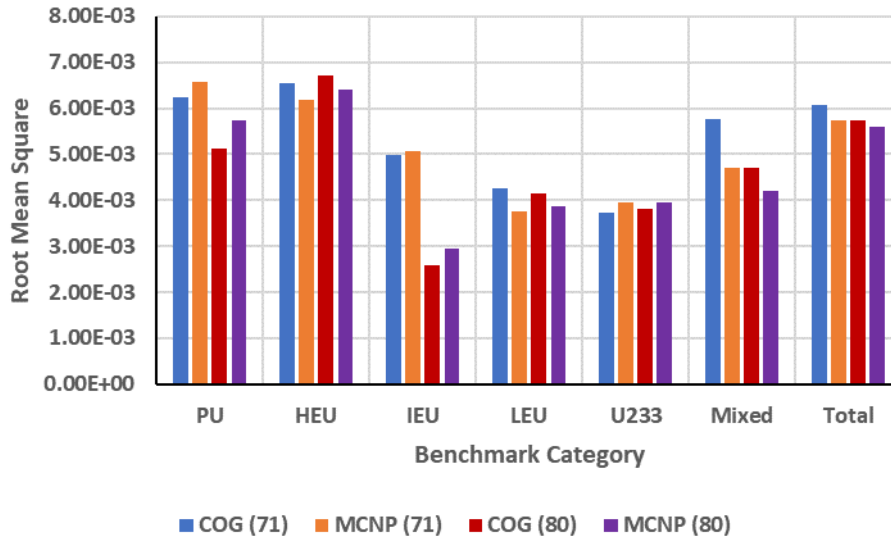


Figure 2. RMSE for 760-Case COG and MCNP Results.

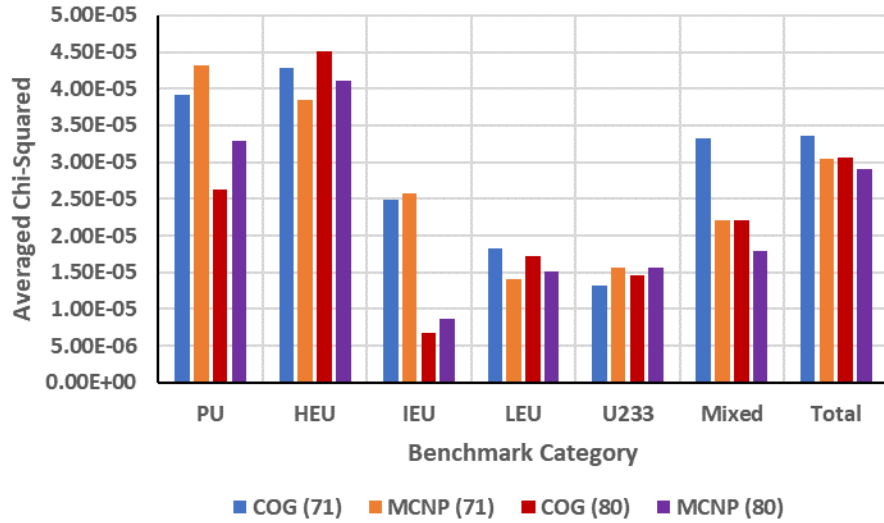


Figure 3. Averaged χ^2 for 760-Case COG and MCNP Results.

3.2 COMPARISON WITH THE ‘MOSTELLER’ 119-CASE VALIDATION SUITE

Selected COG11.1 results were also analyzed using “An Expanded Criticality Validation Suite for MCNP” developed by Mosteller [9], which contains 119 benchmark cases from the ICSBEP Handbook. These cases are summarized in Table 3.

Table 3. Number of Benchmark Cases for 119-Case Comparison.

Category	Number of Cases
Pu	36
HEU	40
IEU	17
LEU	8
²³³ U	18
Total	119

COG11.1 results, identified as “COG11 (LLNL),” were obtained using ENDF/B-VII.1 cross sections and ENDF/B-VII.1 nuclear data processing technique described in Section 2. Additional COG11.1 results, identified as “COG11 (71C),” were obtained using cross sections in the ACE format, which are identical, but re-formatted, to cross sections that are used in MCNP 6.1. The two sets of COG results are compared to each other, and results that were obtained from MCNP 6.1 and MCNP 5 using the same ENDF/B-VII.1 nuclear data. Note that the MCNP 5 results are provided by KAERI [10]. To further compare code performance, COG11.3 was also run using ENDF/B-VIII.0.

The RMSE and the averaged χ^2 for the five different sets of results are compared in Figs. 4 and 5. For ENDF/B-VII.1, COG11 (LLNL) results, shown in blue, are slightly better than MCNP results for the

HEU and ^{233}U cases. The reverse is observed for Pu. The worst performance observed for IEU is due to the same case, MMF008-7, described in the previous subsection.

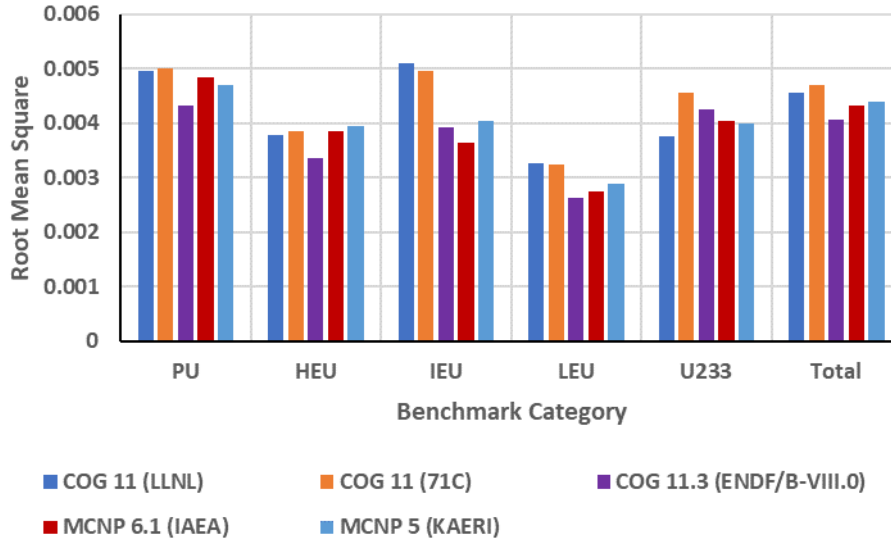


Figure 4. RMSE for 119-Case COG and MCNP Results.

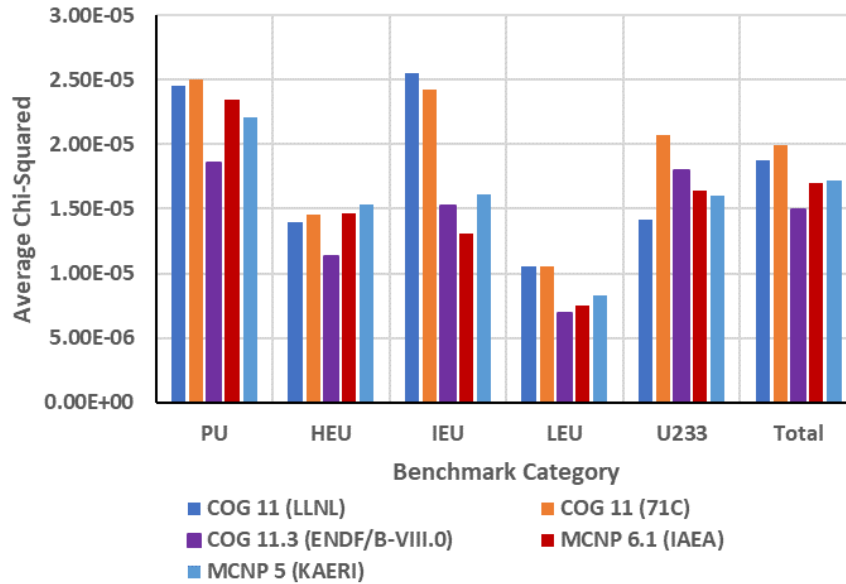


Figure 5. Averaged χ^2 for 119-Case COG and MCNP Results.

Cumulative χ^2 , shown in Fig. 6, compares overall code performances for the 119 cases. All of the cases except COG 11.3 (ENDF/B-VIII.0) cases used ENDF/B-VII.1 data library. Except for the ZEBRA case, good agreement is observed between COG and MCNP results. MCNP results obtained from IAEA and

KAERI produced consistent results. COG11 (LLNL) and COG11 (71C) results also produced consistent results. COG11.3 with ENDF/B-VIII.0 data performed better than COG and MCNP results with ENDF/B-VII.1.

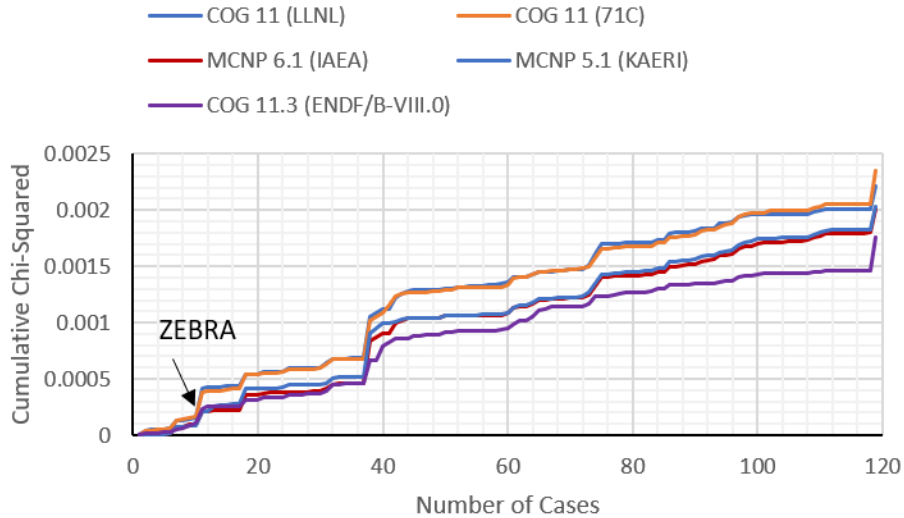


Figure 6. Cumulative χ^2 Comparison for 119-Case COG11 and MCNP Results.

4. CONCLUSIONS

An expanded COG criticality validation suite, consisting of 2,255 benchmark cases, has been established. COG11 results with ENDF/B-VII.1 and ENDF/B-VIII.0 cross section data have been compared with benchmark values from the ICSBEP Handbook. These COG11 results have been also compared with available MCNP results from the 1,081-case ‘Trkov’ and 119-Case ‘Mosteller’ validation suites. Most of the cases agree with the benchmark values within $\pm 3\sigma$. Sources of discrepant results may come from 1) errors in the cross section data, 2) possible errors from the modeling of the benchmark experiments, or 3) errors in the benchmark measurement data itself or its evaluated biases and uncertainties. Outliers beyond the 6σ range of benchmark values are being further examined to identify the cause of such discrepancies. Good agreement was observed between COG11 and MCNP results. COG performed better with the ENDF/B-VIII.0 data library.

An intercomparison project with COG, MCNP, MORET, and SCALE for ENDF/B-VIII.0 and JEFF-3.3 is in progress. The goal of this project is to identify nuclear data processing and benchmark modeling errors through discovery and resolution of discrepant results. At the time of this writing, additional COG cases are being prepared to further expand the validation suite. We anticipate that the final product will consist of more than 2,500 COG cases covering all six categories of the ICSBEP Handbook, which will be beneficial to the international COG user community.

ACKNOWLEDGMENTS

The authors would like to acknowledge Dr. Do H. Kim at KAERI for providing his MCNP results for inter-comparison study. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 and was funded by the U.S. Department of Energy Nuclear Criticality Safety Program.

REFERENCES

1. *COG: A Multiparticle Monte Carlo Code Transport Code, Version 11.1*, CCC-829, Released by Radiation Safety Information Computational Center (2015).
2. *International Handbook of Evaluated Criticality Safety Benchmark Experiments*, NEA/NSC/DOC(95)03, Organization for Economic Cooperation and Development, Nuclear Energy Agency, Paris, France (December 2018 Edition).
3. D. E. Cullen, "POINT 2015 : ENDF/B-VII.1 Final Temperature Dependent Cross Section Library," IAEA-NDS-221 (March 2015).
4. D. Brown, "Use of a continuous integration and deployment software to automate nuclear data V&V," Brookhaven National Laboratory (2014).
5. D. E. Cullen et al., "Thermal Scattering Law Data: Implementation and Testing using the Monte Carlo neutron transport codes COG, MCNP and TART" (May 2003).
6. Andrej Trkov, *On the Benchmarking of New Evaluated Nuclear Data Libraries*, INDC (NDS)-0751, IAEA (March 2018).
7. MCNP6: *Monte Carlo N-Particle Transport Code*, Los Alamos National Laboratory, LA-CP-11-0170 (2012).
8. T. Newton et al., "K-infinity Measurements in ZEBRA Core 8," MIX-MET-FAST-008, *International Handbook of Evaluated Criticality Safety Benchmark Experiments*, NEA/NSC/DOC(95)03, Organization for Economic Cooperation and Development, Nuclear Energy Agency, Paris, France (December 2018 Edition).
9. R. D. Mosteller, "An Expanded Criticality Validation Suite for MCNP," LA-UR-10-06230, Los Alamos National Laboratory (2010)
10. D. H. Kim et al., "Current Status of ACE Format Libraries for MCNP at Nuclear Data Center of KAERI," *Journal of Radiation Protection and Research*, Vol. 41(3) (2016).

APPENDIX A

Table A-1. Number of Subcategorized Benchmark Cases.

Benchmark Category	No. of Cases	Benchmark Category	No. of Cases	Benchmark Category	No. of Cases
pu-comp-inter	1	heu-comp-inter	8	ieu-comp-fast	1
pu-comp-mixed	34	heu-comp-mixed	20	ieu-comp-inter	4
pu-met-fast	99	heu-met-fast	341	ieu-met-therm	82
pu-met-inter	3	heu-met-mixed	8	ieu-met-fast	40
pu-sol-therm	433	heu-met-inter	7	ieu-sol-therm	61
-	-	heu-met-therm	32	-	-
-	-	heu-sol-therm	401	-	-
Total	570	Total	817	Total	188

Table A-1. Number of Subcategorized Benchmark Cases (Continued).

Benchmark Category	No. of Cases	Benchmark Category	No. of Cases	Benchmark Category	No. of Cases
leu-comp-therm	238	u233-comp-therm	5	mix-comp-fast	4
leu-met-therm	48	u233-met-fast	10	mix-comp-inter	1
leu-met-fast	77	u233-sol-inter	33	mix-comp-therm	53
-	-	u233-sol-therm	145	mix-met-fast	44
-	-	-	-	mixed-met-inter	1
-	-	-	-	mix-met-mixed	1
-	-	-	-	mix-sol-therm	20
Total	363	Total	193	Total	124